

P25005

Pattern Recognition Method, Pattern Check Method and Pattern Recognition
Apparatus as well as Pattern Check Apparatus Using the Same Methods

5

Field of the Invention

The present invention relates to a pattern recognition method through which an image or a voice is recognized, a pattern check method and a pattern recognition apparatus as well as a pattern check apparatus using these methods.

10

Background of the Invention

15

Pattern recognition is to classify images taken by cameras or image scanners into a computer. A voice of an object can be picked up through a microphone and taken into a computer. When the object is a human being, his or her facial image or speech voice is to correspond to the object and then classified. Therefore, the pattern recognition technique is summarized to handle the following two variable factors:

20

(1) within-class scatter, i.e., showing different scatters in appearance depending on conditions at recording the object although images or voices are derived from the same object; and

(2) between-class scatter, i.e., showing different patterns due to different objects.

25

In conventional art of the pattern recognition, the following method has been highly assessed: First, assume a sample space corresponding to a set of entire patterns, then apply consistent functions to individual input data so that the within-class scatter can be minimized and the between-class scatter can be maximized, thereby extracting features. The method bases on such a

fundamental model. For instance, Fisher's discriminant analysis (Fukunaga: Introduction to Statistical Pattern Recognition, Academic Press, 1972) has been well known as a typical example, and it has been often used in the fields of character recognition, voice recognition and facial image recognition.

5 However, an assumption of Fisher's discriminant analysis method, i.e., a model of which entire patterns are derived from one distribution, seems sometimes unreasonable viewed from actual problems. For instance, when a system, which checks a face photo on an identification with a facial image taken by a video camera, is considered, this one is made by shooting directly an object and that one is taken indirectly from a printed material. These two materials are compared with each other for determining the identity of the two. However, an assumption that the sets of all the images formed through different processes are derived from one distribution is unreasonable because the images on these two materials differ too much. Actually, it is sometimes difficult for an operator
10 to check a man himself against the photo on his identification.
15

It is thus concluded that Fisher's discriminant analysis method, i.e., entire patterns are described with one distribution and common feature-extraction-functions are consistently applied to input data, which is to be classified, for recognizing patterns, has a limit of accuracy.

20 The inventor of the present invention filed an application (Publication No. EP 0 944 018) with EPO in order to solve this problem. In this patent application, pattern sets A, B are prepared, where pattern set A is an image pattern directly taken a person's face by a video camera while pattern set B is read a photograph of the same person's face by an image scanner. Thus the two
25 sets are formed of the same object but through different processes. From these two pattern sets, (1) distributions of the patterns are found, and (2) perturbation distribution of individual pattern corresponding between pattern sets A and B,

is found, or perturbation distribution of sets of patterns in pattern set A corresponding to each element BI of pattern set B. Then a feature extraction matrix, which minimizes an overlapping cubic volume between the pattern distribution found in (1) and the perturbation distribution found in (2), is found.

5 This feature-extraction-matrix is applied to pattern sets A and B, and each feature amount is calculated. Among the feature amounts, the elements mostly similar to each other are used for determining the identity.

Summary of the Invention

10 The present invention aims to carry out a highly accurate pattern recognition and a pattern check by extracting features which satisfy criteria minimizing the within-class scatter and maximizing the between-class scatter with respect to images formed of the same object but through different processes. Perturbation distributions of individual patterns — corresponding between
15 pattern sets A, B used in the application of EP 0 944 018 — have been broken down into “within-set within-class scatter” and “between-set within-class scatter”. In this context, the class corresponds to a person, and the within-class scatter is a scatter for the same person, and the between-class scatter is a scatter for the different persons.

20 In the present invention, feature extraction matrix AF of pattern set A and feature extraction matrix BF of pattern set B are found so that these matrices AF and BF maximize the between-class scatter with respect to training pattern sets A and B (a training pattern set: a set in pattern sets A and B, actually used for calculation), where the class is not only within a pattern set
25 but also it extends over another pattern set. The pattern sets A, B are obtained based on plural, e.g., two conditions. The scatter here is a scatter between the patterns corresponding to different classes (different objects).

These matrices AF and BF also minimize the within-class scatter with respect to the training pattern sets A and B, where the within-class scatter is a scatter within the classes between patterns A and B corresponding to the same class. These matrices AF and BF are used for calculating feature-amounts of input patterns, and their similarity measures are determined. This method allows a pattern recognition and a pattern check to be superlatively more accurate than conventional methods.

Brief Description of the Drawings

Fig. 1 is a block diagram of a facial image recognition apparatus using a pattern recognition method of the present invention.

Fig. 2 is a flowchart illustrating an operation of off-line recognition in the facial image recognition apparatus shown in Fig. 1.

Fig. 3 is a flowchart illustrating an operation of on-line recognition in the facial image recognition apparatus shown in Fig. 1.

Fig. 4 is a flowchart illustrating an operation of off-line facial image check using a pattern check method of the present invention.

Fig. 5 is a flowchart illustrating an operation of on-line facial image check using a pattern check method of the present invention.

Detailed Description of the Preferred Embodiments

Exemplary embodiments of the present invention are demonstrated hereinafter with reference to the accompanying drawings.

Exemplary Embodiment 1

Fig.1 is a block diagram of a facial image recognition apparatus in which the pattern recognition method of the present invention is applied to identification check. The block diagram of the apparatus in Fig. 1 is realized by

a computer system.

The facial image recognition apparatus comprises the following elements:

(a) video camera 1 or a digital camera (hereinafter referred to as a video camera) for taking a facial image (pattern set A) of a person directly into
5 an electronic recording medium;

(b) image scanner 2 for taking a face photo (pattern set B) on an identification;

(c) image memory 3 for storing an image signal supplied from video camera 1;

10 (d) image memory 4 for storing an image signal supplied from image scanner 2;

(e) memory 5 for storing an operation program;

(f) CPU 6 for performing various calculations and operation controls;

(g) pattern memory 7 for storing the image pattern (hereinafter
15 referred to as pattern A) of pattern set A and a feature pattern calculated by extracting features from pattern set A;

(h) pattern memory 8 for storing the image pattern (hereinafter referred to as pattern B) of pattern set B and a feature pattern calculated by extracting features from pattern set B;

20 (i) a keyboard and display 9 as a system console;

(j) secondary memory device 10 such as a magneto-optic disc for storing large scale data such as image pattern information;

(k) memory 11 for storing feature-extracting matrices (detailed later) and sets of feature-patterns as a data-base (hereinafter referred to as a
25 referential image data-base);

(l) interfaces (I/F) 12 - 16 for transacting data to/from external devices;

(m) system bus 17;

(n) output terminal 18 for outputting a recognition result;

(o) feature extraction matrix memory 19 for storing feature-extraction-matrix AF for pattern A; and

(p) feature extraction matrix memory 20 for storing feature-extraction-matrix BF for pattern B.

The computer system comprises the following elements:

image memories 3, 4;

memory 5;

CPU 6;

10 pattern memories 7, 8;

memory 11;

interfaces (I/Fs) 12 -16;

system bus 17;

output terminal 18; and

15 feature extraction matrix memory 19, 20.

An operation of the apparatus is demonstrated hereinafter. In the facial image recognition apparatus, a large number of identifications have been read by image scanner 2 to produce facial-image-data, which are stored in referential image database 11. A facial image of a person is taken by video camera 1.

20 Then recognize whether the person is registered in database 11 or which identification registered in data-base 11 is mostly similar to the person. The process is roughly divided into two sections. One is an off-line process where feature-extraction-matrix AF for video images and feature-extraction-matrix BF for identification images are calculated, and referential-image database 11 is

25 formed. Another one is an on-line process where a facial image input is determined whether it matches anyone registered in database 11, and when it is determined that the facial image matches someone registered in database 11,

the registered person mostly similar to the facial image is selected from database 11.

In the first place, the off-line process is described with reference to the flowchart shown in Fig. 2. The purpose of the off-line process is to calculate feature extraction matrices AF, BF from training image signals stored in image memories 3 and 4, and to extract a feature vector from the identification images.

The facial images supplied from video camera 1 are stored in image memory 3 via I/F 12 as pattern set A1 (video facial image), and the facial images of identifications taken by image scanner 2 are stored in image memory 4 via I/F 13 as pattern set B1 (identification photo). Respective data are converted into one dimensional data sequence, i.e., pattern A and pattern B, and stored in pattern memories 7, 8 (S11). In this case, generally, video camera 1 produces plural facial images per person, and image scanner produces one facial image from identification images per person.

Feature extraction matrices AF, BF are calculated through steps 12 (S12) – 14 (S14). First, within-class scatter matrix $C_a - w$ and between-class scatter matrix $C_a - b$ of pattern set A1 are calculated following equations 1 and 2. In parallel, between-class scatter matrix $C_b - b$ is calculated following equation 3 (S12).

$$C_a - w = E(i) \{ E(j) \{ (X_{ij}^a - M_i^a) (X_{ij}^a - M_i^a)^T \} \} \dots \dots \dots \text{equa. 1}$$

where, X_{ij}^a represents the "j"th image of the class "i" ("i"th person) in pattern set A1, and M_i^a represents an average image of class "i" in pattern set A1, $E(i) \{ \cdot \}$ represents an operator which averages $\{ \cdot \}$ with respect to $\{ i \}$, and T shows transposition of the matrix and vector.

$$C_a - b = E(i) \{ (M_i^a - M^a) (M_i^a - M^a)^T \} \dots \dots \dots \text{equa. 2}$$

where, M^a represents an average image of entire pattern set A1.

$$C_b - b = E(i) \{ (M_i^a - M^a) (M_i^a - M^a)^T \} \dots\dots\dots \text{equa. 3}$$

where, M^a represents an average image of entire pattern set B1.

Next, equation 4 calculates a cross-correlation matrix $C_{ab} - r$ of pattern set A1 and pattern set B1 (S13).

$$C_{ab} - r = E(i) \{ (M_i^a) (M_i^b)^T \} \dots\dots\dots \text{equa. 4}$$

Then equations 5, 6 calculate feature-extraction-matrices AF, BF with respect to pattern sets A1, B1 (S14).

$$AF = (P1w)^T C_a - w^{-1/2} \dots\dots\dots \text{equa. 5}$$

where, P1w represents a matrix of eigenvector corresponding to "m" pcs. of maximum eigenvalues of matrix $(C_a - w^{-1/2}) C_a - b (C_a - w^{-1/2})$, where $K^{(-1/2)}$ represents a positive semi-definite square root matrix with respect to matrix K.

$$BF = Q P2b^T C_b - b^{-1/2} \dots\dots\dots \text{equa. 6}$$

where, matrix P2b represents an orthogonal matrix obtained by singular value decomposition of matrix $C_0 = C_a - b^{(-1/2)} C_{ab} - r C_a - b^{(-1/2)}$, and satisfies $C_0(P2b) = (P1b)D$. "D" represents a diagonal matrix, P1b shows an orthogonal matrix, matrix $Q = (P1w)^T (C_a - w)^{(-1/2)} R^*$, where R^* is a quasi-inverse matrix of matrix $R = (P2b)^T (C_a - b)^{(-1/2)}$.

The validity of this feature-extraction function is described as follows: First, between $C_{ab} - r$ and matrix $C_{ab} - w$, a relation shown by equation 8 exists. Within-class scatter matrix $C_{ab} - w$ shown by equation 7 extends over pattern set A1 and pattern set B1. In this connection, the minimization of $C_{ab} - w$ is equivalent to the minimization of $C_a - b$ and $C_b - b$, as well as the maximization of $C_{ab} - r$.

$$C_{ab} - w = E(i) \{ (M_i^a - M_i^b) (M_i^a - M_i^b)^T \} \dots\dots\dots \text{equa. 7}$$

$$C_{ab} - w = C_a - b + C_b - b - 2 C_{ab} - r \dots\dots\dots \text{equa. 8}$$

Assessment standard S shown by equation 9 minimizes the within-class scatter of respective pattern sets A1, B1, and is maximized when it maximizes

the between-class scatter only when the scatter extends over two pattern sets A1, B1.

$$S = \text{tr} [Ca - w^{-1} Ca - b] - 2 \text{tr} [(Ca - b + Cb - b)^{-1} Cab - w] \dots \text{equa. 9}$$

Feature extraction functions $fa = AF(X^a)$, $fb = BF(X^b)$, which maximize
5 assessment standard S , agree with feature extraction matrices AF , BF which maximize equation 10.

$$S = \text{tr} [Ca - w^{(-1)} Ca - b] + 2 \text{tr} [(Ca - b + Cb - b)^{-1} Cab - r] - I \dots \text{equa. 10}$$

where, "I" represents a unit matrix.

One of quasi-optimum feature-extraction-matrices is AF or BF shown by equa. 5

10 or 6. These matrices maximize equation 10.

Then in order to build a referential image database 11, plural facial images of identifications are taken by image scanner 2. Feature vector "fB1i" is calculated using matrix BF and following equation 11 with respect to typical pattern B1i of respective classes (respective persons), and the vector is
15 registered to database 11 (S15).

$$fB1i = FB (B1i) \dots \dots \dots \text{equa. 11}$$

The descriptions above have discussed the off-line process.

The on-line process is now demonstrated hereinafter with reference to the flowchart shown in Fig. 3. The purpose of the on-line process is to determine
20 whether a facial image input matches a person registered in referential image database 11, and to select the most similar one from the database when it matches one of registered persons.

A facial image directly taken by video camera 1 is stored in image memory 3, then pattern "a" input is converted into one-dimensional data "pattern A2j",
25 which then transferred to pattern memory 7 (S30). Feature-extraction-matrix AF , found in the off-line process and stored in memory 19, is applied to pattern A2j stored in memory 7, then feature vector $fA2j$ is calculated following equation

12 (S31).

$$fA2j = AF(A2j) \dots\dots\dots \text{equa. 12}$$

Next, vary entry-index "i" of database 11 so that a feature vector similar to fA2j is selected from database 11, and an optimum matching process is carried out (S32). At least one result of the process is output as a recognition result to output terminal 18 (S33).

In the description discussed above, referential image database 11 is formed by collecting facial images from identifications through image scanner 2; however, database 11 can be formed by inputting facial images through video camera 1. Image scanner 2 and video camera 1 are used as pattern input means; however, either one or another input means is acceptable.

Facial image data, pattern sets A1, B1 and database 11 are produced in the off-line process; however, those materials may have been stored in secondary memory device 10 so that they can be retrieved from memory device 10.

In this embodiment, the facial image recognition apparatus is used as an example; however, image data of cars, assembled parts, or voice data, character data can be converted into pattern data, so that the apparatus can be used as a pattern recognition apparatus, and various applications can be expected.

20 Exemplary Embodiment 2

A facial image checking apparatus, in which a pattern check method is applied to identification check, is demonstrated hereinafter. The apparatus in accordance with the second embodiment has the same block diagram as that shown in Fig. 1. Thus the block diagram and the description thereof are omitted here.

The facial image checking apparatus is used for determining whether or not a facial image taken from an identification is identical with another facial

image taken by a video camera. The determination process is divided into two sections. One is an off-line process comprising the steps of:

(a) obtaining facial images from a lot of identifications through image scanner 2;

5 (b) obtaining a facial image by shooting a person's face with video camera 1; and

(c) calculating feature-extraction-matrices AF, BF from these facial images.

Another section is an on-line process comprising the steps of:

10 (d) calculating a feature amount using matrices AF, BF from the facial image (pattern A) supplied from image scanner 2 and the facial image (pattern B) supplied from video camera 1; and

(e) determining whether or not both the facial images are identical based on similarity measure of the feature amounts.

15 In the first place, the off-line process is demonstrated with reference to the flowchart shown in Fig. 4. The purpose of the off-line process is to calculate feature-extract-matrices AF, BF.

First, an image signal of a person's face shot by video camera 1 is stored in image memory 3 via I/F 12. At the same time, a facial image is taken from
20 the person's identification by image scanner 2 and stored in image memory 4 via I/F 13 (S40). This procedure is repeated until a necessary number of images enough for feature extraction matrix F to learn is collected (e.g. 15,000 persons).

The facial images taken by video camera 1 are stored in memory 3 as pattern set A1 (facial image), and the facial images of the identifications taken
25 by image scanner 2 are stored in memory 4 as pattern set B1 (a photo on an identification). Respective pattern sets are converted to one-dimensional data sequence, i.e., pattern A1 and pattern B1, and they are stored in pattern

memories 7, 8 (S41).

Feature-extraction-matrices AF , BF are calculated following the steps of S42 – S44. First, within-class scatter matrix $Ca - w$ and between-class scatter matrix $Ca - b$ of pattern set $A1$ are calculated following equations 1, 2. At the same time, within-class scatter matrix $Cb - b$ of pattern set $B1$ is calculated following equation 3 (S42). Then cross-correlation matrix $Cab - r$ of pattern sets A, B is calculated following equation 4 (S43).

Next, feature-extraction-matrices AF , BF are calculated following equations 5, 6 with respect to pattern sets $A1$, $B1$. The validity of these feature-extraction-functions are described as follows: First, between $Cab - r$ and matrix $Cab - w$, a relation shown by equation 8 exists. Within-class scatter matrix $Cab - w$ shown by equation 7 extends over pattern set $A1$ and pattern set $B1$. In this connection, the minimization of $Cab - w$ is equivalent to the minimization of $Ca - b$ and $Cb - b$, as well as the maximization of $Cab - r$.

Assessment standard S shown by equation 9 minimizes the within-class scatter of respective pattern sets $A1$, $B1$, and is maximized when it maximizes the between-class scatter only when the scatter extends over two pattern sets $A1$, $B1$.

Feature extraction functions $fa = AF(X^a)$, $fb = BF(X^b)$, which maximize assessment standard S , agree with feature extraction matrices AF , BF which maximize equation 10. One of quasi-optimum feature-extraction-matrices is AF or BF shown by equa. 5 or 6. These matrices maximize equation 10. The descriptions above have discussed the off-line process.

Next, the on-line process is demonstrated with reference to Fig. 5. The purpose of the off-line process is to determine whether or not the facial image of an identification supplied from image scanner 2 is identical with the facial image taken by video camera 1.

The facial image shot with video camera 1 undergoes A/D conversion, and is stored in image memory 3, and the facial image taken by image scanner 2 is stored in image memory 4. These facial images are retrieved from the memory, and converted to, e.g., one-dimensional data sequences, then they are stored in pattern memories 7, 8 (S50).

Next, feature-extraction-matrices AF, BF found in the off-line process and stored in memory 19, 20 are applied to patterns "a2", "b2" stored in memory 7, 8, then feature vector "fA2", "fB2" are found using equations 13, 14 (S51).

$$fA2 = AF (a2) \dots\dots\dots \text{equa. 13}$$

$$fB2 = BF (b2) \dots\dots\dots \text{equa. 14}$$

Determine, then, whether or not these facial images are identical based on the similarity measure of feature vectors fA2, fB2. In other words, determine whether these facial images are derived from the same class (S52). The determination result (Y/N) is output to output terminal 18 for a pattern check (S53).

In the descriptions discussed above, facial image data, pattern sets A1, B1 are obtained in the off-line process; however, they have been stored in secondary memory device 10 and can be retrieved therefrom. Further, those data obtained in the off-line process can be stored in secondary memory device 10 and they can be updated.

On the premise that there are different distributions between pattern sets to be compared with, the present invention carries out an optimum feature extraction which satisfies a consistent standard, i.e., within-class scatter extending over two distributions is minimized and between-class scatter is maximized with respect to the two distributions respectively. As a result, a pattern recognition with higher accuracy than that through the conventional method, as well as a pattern check, can be realized.